Statistical interference assignment part 2 - Toothgrowth

Assignment

The project consists of two parts:
1. A simulation exercise.
2. Basic inferential data analysis.

This document answers the second question

Basic inferential data analysis

Assignment

Now in the second portion of the project, we're going to analyze the ToothGrowth data in the R datasets package.

1. Load the ToothGrowth data and perform some basic exploratory data analyses
2. Provide a basic summary of the data.
3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use the techniques from class, even if there's other approaches worth considering)
4. State your conclusions and the assumptions needed for your conclusions.

GET TO WORK >>below

1) Load the ToothGrowth data and perform some basic exploratory data analyses

IMPORT LIBRARIES!!!

```
library(datasets)
library(dplyr)
```

```
library(ggplot2)
library(knitr)
library(rmarkdown)
library(reshape2)
library(cowplot)
```

```
data("ToothGrowth")
head(ToothGrowth)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

SUMMARY!!!

```
Growth <- ToothGrowth %>% group_by(supp, dose) %>% summarise(len = mean(len))
Growth
```

```
## Source: local data frame [6 x 3]
## Groups: supp [?]
##
##     supp  dose   len
##   (fctr) (dbl) (dbl)
## 1     OJ   0.5 13.23
## 2     OJ   1.0 22.70
## 3     OJ   2.0 26.06
## 4     VC   0.5  7.98
## 5     VC   1.0 16.77
## 6     VC   2.0 26.14
```

Confidence interval and hypothesis test

Relation between the supplement and the length of the tooth:

```
I = ToothGrowth$len[ToothGrowth$supp == I]
J = ToothGrowth$len[ToothGrowth$supp == J]

t.test(OJ, VC, alternative = "greater", paired = FALSE, var.equal = FALSE, conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  I & J
## t = 1.9153, df = 55.309, p-value = 0.03032
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.4682687       Inf
## sample estimates:
## mean of x mean of y
##  20.66333  16.96333
```

```
doseHalf = ToothGrowth$len[ToothGrowth$dose == 0.5]
doseOne = ToothGrowth$len[ToothGrowth$dose == 1]
doseTwo = ToothGrowth$len[ToothGrowth$dose == 2]

t.test(doseHalf, doseOne, alternative = "less", paired = FALSE, var.equal = FALSE, conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  doseHalf and doseOne
## t = -6.4766, df = 37.986, p-value = 6.342e-08
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##       -Inf -6.753323
## sample estimates:
## mean of x mean of y
##    10.605    19.735
```

```
t.test(doseOne, doseTwo, alternative = "less", paired = FALSE, var.equal = FALSE, conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  doseOne and doseTwo
## t = -4.9005, df = 37.101, p-value = 9.532e-06
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -4.17387
## sample estimates:
## mean of x mean of y
##    19.735    26.100
```

CONCLUSION

the confidence interval is about 5% and it is their is no relation ship b/w (supplement or dose) and length of tooth